Simulate Collusion: Application Examples for Cartel Simulation based on stochastic interest rates

Set knitr Options

```
knitr::opts_chunk$set(echo = TRUE, warning = FALSE, tidy=TRUE, tidy.opts = list(width.cutoff=
80))
```

Clear Environment

```
rm(list = ls())
```

Load Packages

Set Seed for Reproducibility

```
set.seed(123)
```

Read in Data

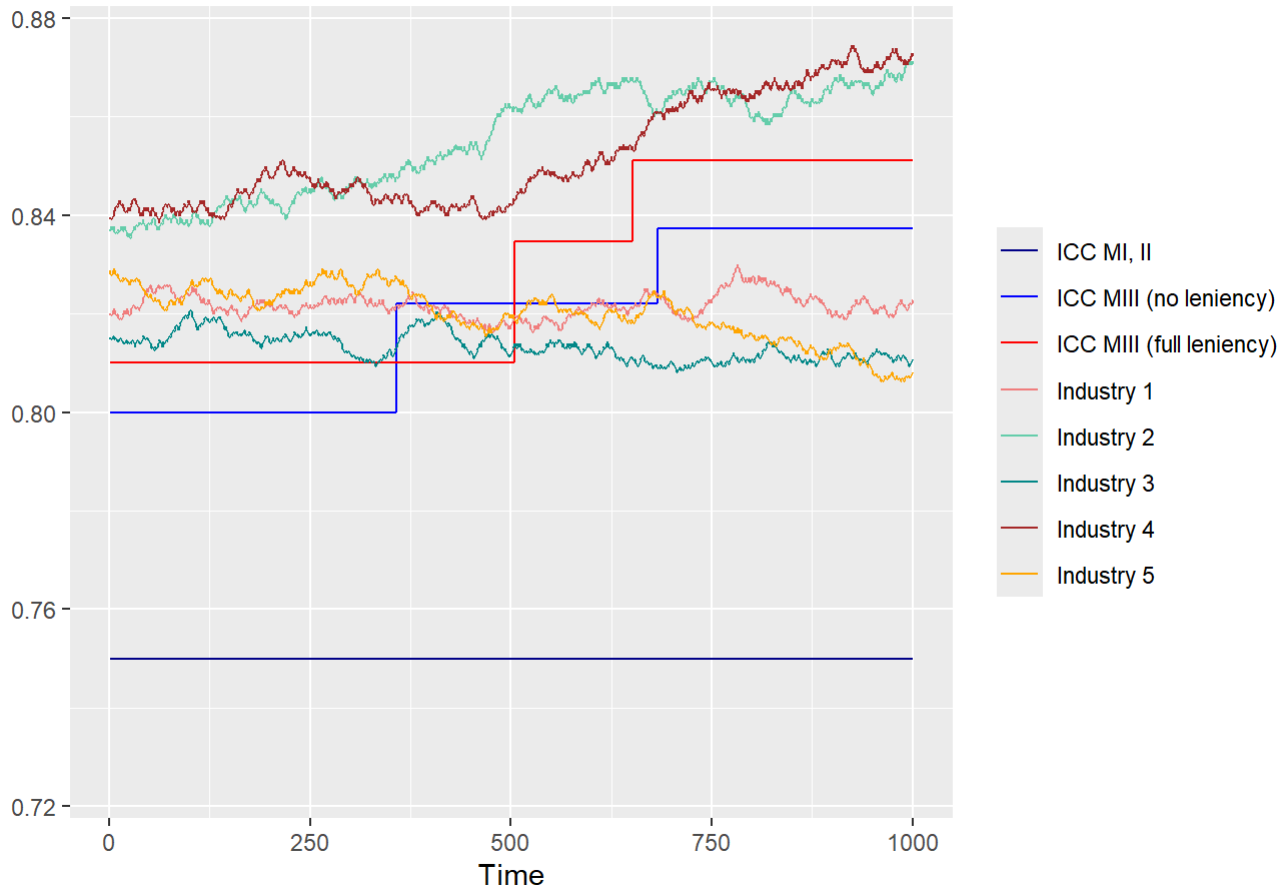
```
# Baseline Model 1 with default parameters.
sim_list <- sim_col_r()

# Model I in Bellert, Günster (2025) sim_list <- sim_col_r(model = 1, periods =
# 1000, n_industries = 300, r_min = 0.001, r_max = 0.3, n_firms_min = 2,
# n_firms_max = 100, alpha = 500)
cartels_duration <- get_durations(sim_list$cartels_detected, sim_list$cartels_undetected,
    sim_list$interest_r, sim_list$deltas, sim_list$parms, model = 1)
cartels_duration <- data.frame(cartels_duration)
```

Add Nonlinear Interdependencies between Variables

```
cartels_duration <- add_nonlinears(cartels_duration, model = 1)
```

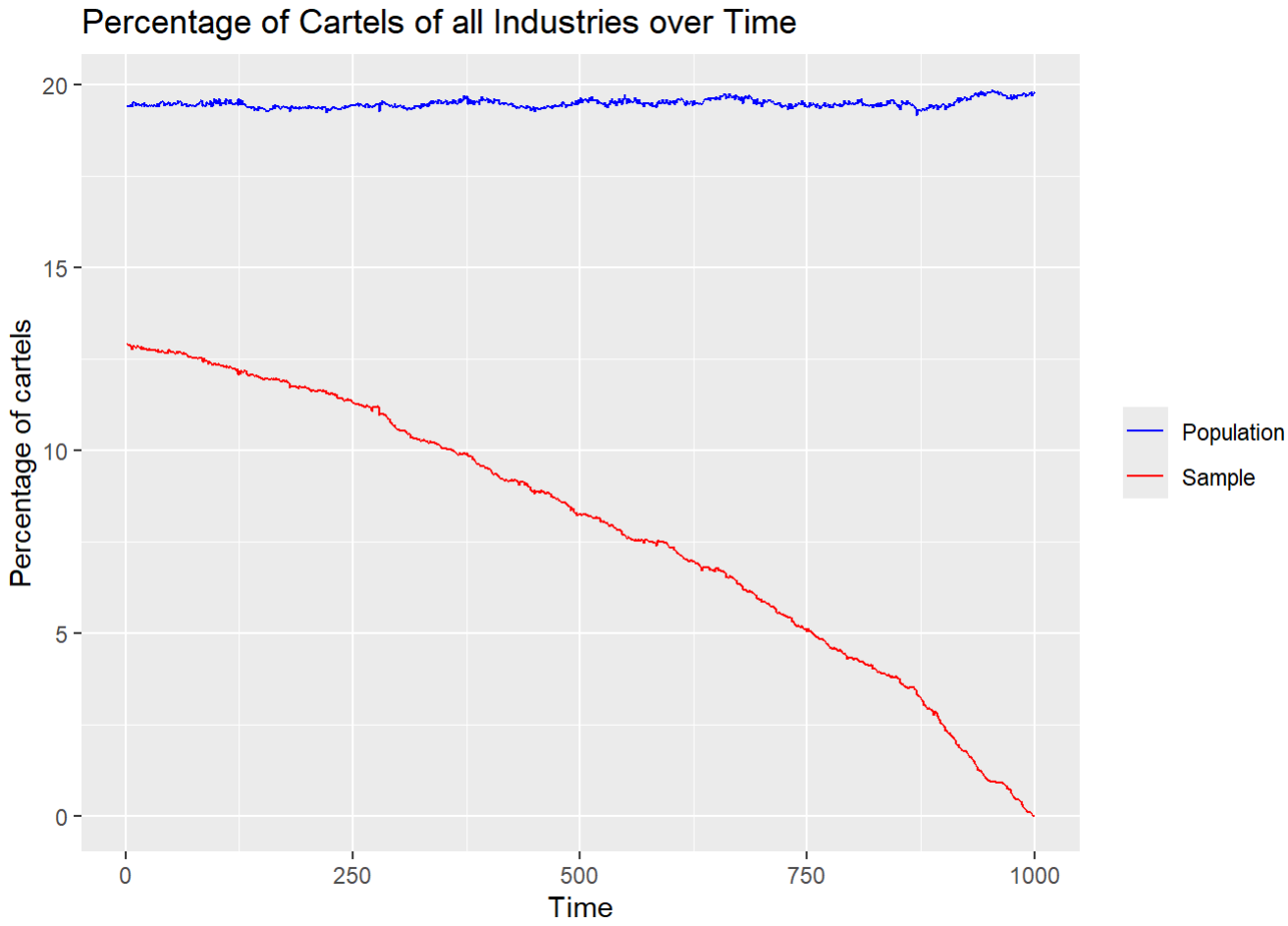
Plot for ICC and Discount Factor of 5 Example Industries

```
set.seed(123)
n_firms <- 4
gamma <- 0.8
theta <- 1
sigma_all <- 0.25
sigma_t <- 1 - (1 - sigma_all)^(1/200)
struc <- 1

r_min <- 0.001
r_max <- 0.3

sim_list_M1 <- sim_col_r(model = 1, n_firms_min = 4, n_firms_max = 4, n_industries = 5,
    sigma = 0.25, gamma = 0.8, theta = 1, struc = 1)
sim_list_M3_1 <- sim_col_r(model = 3, n_firms_min = 4, n_firms_max = 4, n_industries = 5,
    sigma = 0.25, gamma = 0.8, theta = 1, struc = 1)
sim_list_M3_0 <- sim_col_r(model = 3, n_firms_min = 4, n_firms_max = 4, n_industries = 5,
    sigma = 0.25, gamma = 0.8, theta = 0, struc = 1)
x <- sim_list_M3_1$ICC
x <- data.frame(x)
df <- cbind(sim_list_M1$ICC[, 1, 1], sim_list_M3_1$ICC[, 1, 1], sim_list_M3_0$ICC[,
    1, 1], sim_list_M3_1$deltas[, , 1])
df <- data.frame(df)
sim <- ts(df)
colnames(sim) <- c("ICC MI, II", "ICC MIII (no leniency)", "ICC MIII (full leniency)",
    paste("Industry", 1:ncol(sim_list_M3_0$deltas[, , 1])))
pallete <- c("blue4", "blue", "red", "lightcoral", "aquamarine3", "cyan4", "brown",
    "orange")
y_axis <- c(0.725, 0.875)  # for singular picture
autoplot(sim) + ylim(y_axis) + xlab("Time") + ylab("") + scale_colour_manual(values = pallet
e) +
    theme(legend.title = element_blank())
```



Plot Cartel Time Series

```
c_det <- rowSums(sim_list$cartels_detected)/(dim(sim_list$cartels_detected)[2] *
    dim(sim_list$cartels_detected)[3])  # number of detected cartels / number of industries
c_pop <- rowSums(sim_list$cartels_population)/(dim(sim_list$cartels_population)[2] *
    dim(sim_list$cartels_population)[3])  # number of cartels / number of industries
sim_cartels <- ts(data = cbind(c_pop, c_det))
colnames(sim_cartels) <- c("Population", "Sample")
pallete <- c("blue", "red")

autoplot(sim_cartels * 100) + xlab("Time") + ylab("Percentage of cartels") + ggtitle("Percent
age of Cartels of all Industries over Time") +
    scale_colour_manual(values = pallete) + theme(legend.title = element_blank())
```



Summary Statistics

```
describe(cartels_duration)
```

	v...	n	mean	sd	median	trimmed	
	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	
parm_id	1	2708	1.784749e+01	8.680428e+00	1.900000e+01	1.801937e+01	8.895600
industry	2	2708	1.453877e+01	8.668694e+00	1.400000e+01	1.435517e+01	1.037820
cartel	3	2708	5.522526e+00	6.464178e+00	3.000000e+00	4.086716e+00	2.965200
detected	4	2708	3.212703e-01	4.670506e-01	0.000000e+00	2.767528e-01	0.000000
nTc	5	2708	8.822009e-01	1.142817e+00	0.000000e+00	6.692804e-01	0.000000
rep_off	6	2708	2.407681e-01	4.276287e-01	0.000000e+00	1.761993e-01	0.000000
start	7	2708	4.002441e+02	3.407651e+02	3.700000e+02	3.824267e+02	5.233578
end	8	2708	6.455155e+02	3.247160e+02	7.130000e+02	6.725992e+02	4.255062
duration	9	2708	2.462714e+02	3.218551e+02	7.700000e+01	1.859110e+02	1.126776
lduration	10	2708	3.858064e+00	2.300323e+00	4.356709e+00	3.875425e+00	3.045444

1-10 of 17 rows | 1-8 of 14 columns Previous **1** 2 Next

Correlations

```
df_cor <- Filter(function(x) sd(x) != 0, cartels_duration)
df_cor <- select(df_cor, -c(parm_id, industry, cartel, lduration, nfirms_sigma, mean_delta))
cor(df_cor)
```

```
##             detected         nTc     rep_off       start          end
## detected  1.00000000  0.51595326  0.36350774 -0.41945505 -0.303372079
## nTc       0.51595326  1.00000000  0.85175975  0.07907463  0.184004536
## rep_off   0.36350774  0.85175975  1.00000000  0.17379382  0.205297363
## start    -0.41945505  0.07907463  0.17379382  1.00000000  0.533071039
## end      -0.30337208  0.18400454  0.20529736  0.53307104  1.000000000
## duration  0.13803060  0.10191962  0.02311749 -0.52094361  0.444498284
## n_firms  -0.43763128 -0.49004536 -0.36400202  0.29965026 -0.172823720
## sigma     0.15271088  0.23893317  0.24120670  0.07964984 -0.071565878
## mean_r    0.06544833  0.07250863  0.07290290 -0.04677827  0.008801452
## var_r     0.09799320  0.06707937  0.01295915 -0.33498134  0.297613541
##             duration      n_firms        sigma       mean_r       var_r
## detected  0.13803060 -0.437631281  0.152710884  0.065448333  0.09799320
## nTc       0.10191962 -0.490045364  0.238933168  0.072508629  0.06707937
## rep_off   0.02311749 -0.364002016  0.241206701  0.072902899  0.01295915
## start    -0.52094361  0.299650255  0.079649844 -0.046778271 -0.33498134
## end       0.44449828 -0.172823720 -0.071565878  0.008801452  0.29761354
## duration  1.00000000 -0.491615547 -0.156531533  0.058406324  0.65492150
## n_firms  -0.49161555  1.000000000 -0.009611012 -0.377181218 -0.32922171
## sigma    -0.15653153 -0.009611012  1.000000000 -0.017069398 -0.08019530
## mean_r    0.05840632 -0.377181218 -0.017069398  1.000000000  0.03235240
## var_r     0.65492150 -0.329221714 -0.080195301  0.032352398  1.00000000
```

Mean Comparison Tests for Sumstats

```
t.test(duration ~ detected, data = cartels_duration, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  duration by detected
## t = -8.0336, df = 2219.6, p-value = 1.521e-15
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal
to 0
## 95 percent confidence interval:
##  -118.33914  -71.90085
## sample estimates:
## mean in group 0 mean in group 1
##        215.7122        310.8322
```

```
t.test(n_firms ~ detected, data = cartels_duration, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  n_firms by detected
## t = 26.608, df = 1938.1, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal
to 0
## 95 percent confidence interval:
##  1.233705 1.430043
## sample estimates:
## mean in group 0 mean in group 1
##        4.760609        3.428736
```

```
t.test(sigma ~ detected, data = cartels_duration, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  sigma by detected
## t = -8.3003, df = 1851.6, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal
to 0
## 95 percent confidence interval:
##  -0.03438520 -0.02124136
## sample estimates:
## mean in group 0 mean in group 1
##       0.2336235       0.2614368
```

Linear Regression

```
reg <- lm(lduration ~ start + n_firms + sigma + mean_r + var_r, data = cartels_duration)
summary(reg)
```

```
##
## Call:
## lm(formula = lduration ~ start + n_firms + sigma + mean_r + var_r,
##     data = cartels_duration)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.6052 -1.2476 -0.0461  1.1861  3.6685
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.188e+00  1.597e-01  51.272   <2e-16 ***
## start       -1.348e-03  9.248e-05 -14.580   <2e-16 ***
## n_firms     -8.636e-01  2.398e-02 -36.014   <2e-16 ***
## sigma        1.270e-01  3.422e-01   0.371    0.711
## mean_r      -3.428e+00  3.367e-01 -10.181   <2e-16 ***
## var_r        1.980e+03  9.420e+01  21.022   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.504 on 2702 degrees of freedom
## Multiple R-squared:  0.5732, Adjusted R-squared:  0.5724
## F-statistic: 725.7 on 5 and 2702 DF,  p-value: < 2.2e-16
```

Hazard Rate Estimation

```
periods <- dim(sim_list$cartels_population)[1]
data_pop <- cartels_duration
data_sample <- data_pop[data_pop$detected == 1, ]
data_undetect <- data_pop[data_pop$detected == 0, ]

data_pop$dead <- ifelse(data_pop$end < periods, 1, 0)
data_sample$dead <- ifelse(data_sample$end < periods, 1, 0)
data_undetect$dead <- ifelse(data_undetect$end < periods, 1, 0)

hazC <- phreg(Surv(duration, dead) ~ start + n_firms + sigma + mean_r + var_r, data = data_po
p,
    dist = "weibull")
hazS <- phreg(Surv(duration, detected) ~ start + n_firms + sigma + mean_r + var_r,
    data = data_sample, dist = "weibull")
hazU <- phreg(Surv(duration, dead) ~ start + n_firms + sigma + mean_r + var_r, data = data_un
detect,
    dist = "weibull")

stargazer(hazS, hazU, hazC, title = "HR Regression for Cartel Duration on Model I",
    type = "text", column.labels = c("HRSample", "HRUndetect", "HRCartels"), df = FALSE,
    digits = 3)
```

```
##
## HR Regression for Cartel Duration on Model I
## ============================================================
##                            Dependent variable:
##                -----------------------------------------
##                  duration              duration
##                  HRSample      HRUndetect     HRCartels
##                     (1)           (2)            (3)
## ------------------------------------------------------------
## start             0.002***      -0.0004***      0.00004
##                   (0.0001)       (0.0001)      (0.0001)
##
## n_firms             0.030         2.425***      0.578***
##                    (0.031)       (0.183)        (0.024)
##
## sigma              1.110**       -0.795**        0.357
##                    (0.440)       (0.346)        (0.266)
##
## mean_r             -0.012        18.298***      2.774***
##                    (0.373)       (1.438)        (0.280)
##
## var_r           -1,829.672*** -38,029.590*** -4,862.372***
##                   (185.539)    (2,337.278)     (227.926)
##
## log(scale)        5.904***      21.292***      8.799***
##                    (0.136)       (1.582)        (0.265)
##
## log(shape)        0.294***      -0.275***      -0.522***
##                    (0.028)       (0.022)        (0.017)
##
## ------------------------------------------------------------
## Observations         870           1,838          2,708
## Log Likelihood  -5,686.590     -3,655.444    -10,784.150
## ============================================================
## Note:                          *p<0.1; **p<0.05; ***p<0.01
```

Future Research: Create Paneldata

```
df_panel <- get_paneldata(sim_list$cartels_population, sim_list$cartels_detected,
    sim_list$parms, sim_list$interest_r, sim_list$sigmas)
```